Response to comments from anonymous referee #4

1. *There is a strong need for field-based measurements of emission factors and emission rates from road traffic to provide data collected under realistic conditions as opposed to the idealized driving cycles used for regulatory purposes. The authors make a compelling case for on-road measurements of the kind which they describe and their methods appear to provide a useful advance.*
   The paper is extremely short on experimental detail and more is needed in the main paper to provide value to the readers. Issues which need to be addressed include quality assurance, the measurement technique for PAH (not currently mentioned at all) and the location(s) of the off-road background measurements. Was a single site used and is this representative for all of the freeways sampled? Regarding the PAH, measurements were made with a continuous PAS sensor, and there need to be caveats over the considerable sensitivity variations for different PAH compounds and the fact that changes in the PAH mixture may manifest themselves as apparent changes in concentration.

Roadway concentrations are elevated compared to off-road or ambient concentrations, and these concentrations act as a kind of baseline concentration to which emissions from current vehicles are superimposed. Our goal was to measure the increment in pollutant concentration over and above this baseline. Background concentration was calculated as the first percentile of the concentrations on each freeway segment for each run. This percentile approach also gave us a baseline measure that was the spatial and temporal equivalent of elevated roadway concentration measurements, and directly comparable to the traffic emissions we were measuring since they were made with the same instruments. The percentile approach was briefly mentioned in the manuscript at Lines 11-13/Pg. 18722, and we have provided more details in the revised version. Also, for clarity “background” has been replaced by “baseline.” The relevant section now reads as:

> “Roadway baseline values in Equation 1 were estimated as the first percentile of pollutant concentrations observed on each freeway link. Since we were attempting to measure concentration increases from current traffic during a short time interval—over and above the elevated concentrations already present on the freeways—a percentile concentration value was the most appropriate indicator of baseline roadway concentrations. The lowest few concentration percentiles were relatively insensitive to superimposed traffic (CO_2 percentile profiles for a subset of runs have been plotted in Figure S3 in SI). Concentrations at a distant location away from the freeway were lower and would have provided an artificially low baseline estimate and upwardly biased the EFs.”

PAH measurement technique: We have added more detail on PAH measurement technique and acknowledged the lack of PAH chemical composition. The revised text reads as:

> "Polycyclic aromatic hydrocarbons (PAH) were measured using PAS 2000 unit (EcoChem Analytics, League City, TX), which provides a mass concentration of particle-bound PAH species (PB-PAH). The instrument's response depends on physical characteristics of the particulate (size and shape) as well as the chemical composition of the PAHs. For example, benzo[a]pyrene with five rings produces a stronger photoemissions signal compared with particles coated with an equal mass of chrysene with four rings (Niessner et al., 1986). However, continuous PAH measurements using this instrument have been used a "semi-quantitative" measure in many roadway studies (For example, Marr et al., 2004). In these studies and ours, researchers have observed strong correlations between PB-PAH, BC and PN concentration (see Figure S2 in the SI), indicating that PAH instrument response was approximately proportional to primary vehicle emissions over a magnitude change in concentration of BC and PN."
>
> "

We also added figure S-2 to the Supplement that shows good correlations between the PAS2000 response versus PN or BC by freeway segment.

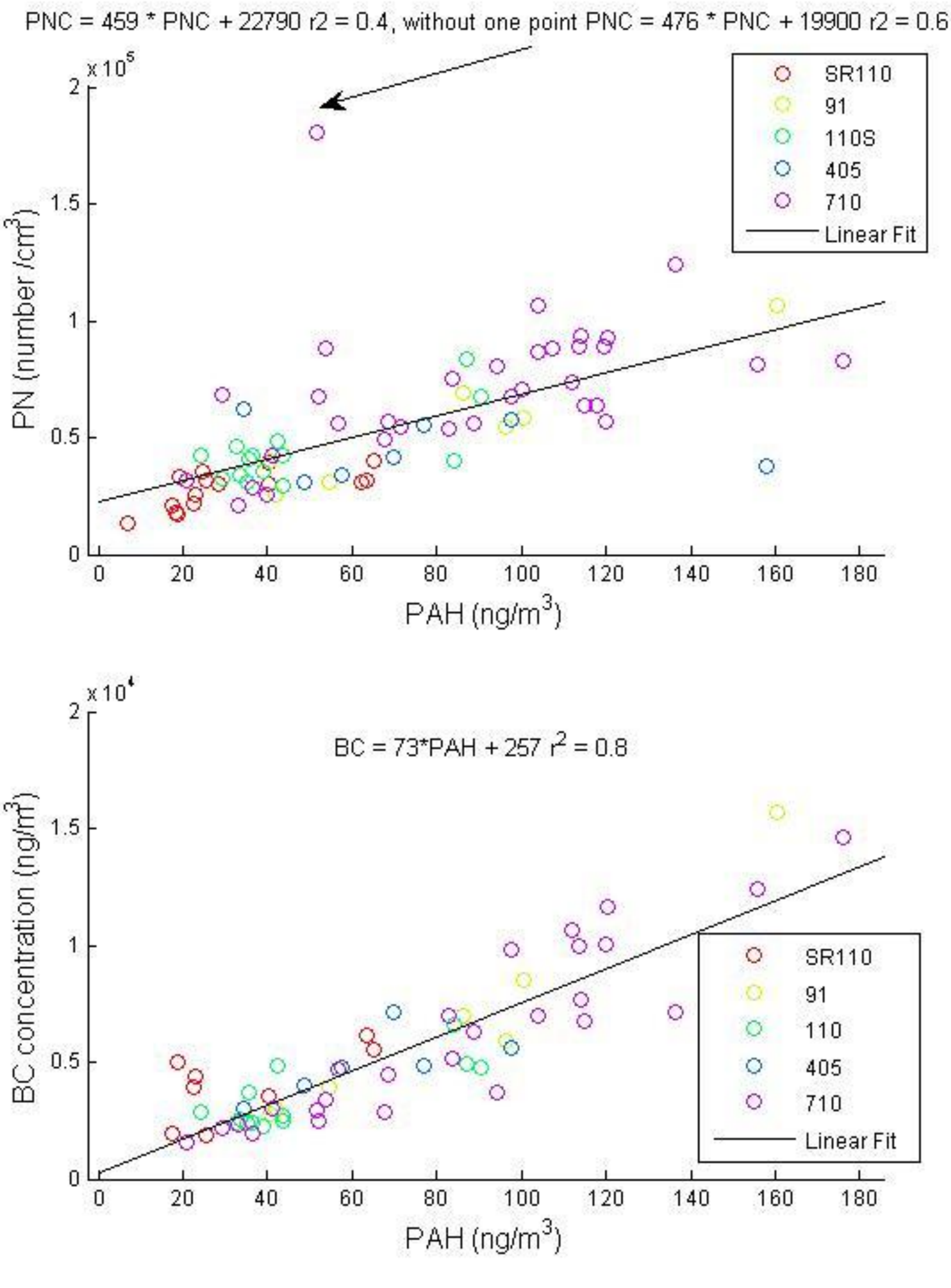

Figure S2: PAH correlation with BC and PNC.

2. *There is one very major issue which it is essential to address in depth before the work should be accepted for final publication. The abstract states that "...... it captured much or most of the variability in EFs due to inter-vehicle differences". It is very unclear to the Referee to what extent inter-vehicle differences are captured.*

   We recognize that use of the term "inter-vehicle" wasn't unambiguous so changed the term to "intra-fleet" variability. Our study set out to develop an efficient methodology to measure EFs, which also captures the variation in EFs resulting from "a wide range of real-driving conditions" on different roadways (i.e., multiple freeways). The variability in our results thus reflects the **total** variation in EFs due to differences in fleet composition, driving conditions and roadway conditions such as grade, i.e., the totality of factors that will influence emissions. Each run reflected a different fleet composition not only in terms of fraction that was HDD, but also a new set of vehicles under a new set of driving conditions.

3. *It is indicated in the methods section that 10 second averages were used to determine concentrations on freeway segments. Does this mean that the data in the histograms in Figures 2 and 3 are derived from 10 second averages? If this is the case, why is the total number of observations so low? Is a 10 second observation commensurate with determining the emission factor from an individual vehicle? This seems very unlikely when the sampling vehicle is travelling on a busy highway and is potentially influenced by the plumes of many vehicles. It is essential that this point should be resolved or else the meaning of the distributions shown in the figures will be entirely obscure. The authors put interpretation on the spread of EF values which may not be warranted.*

   10 second averages were not used to estimate EFs. EFs were calculated using freeway run segment medians. We have revised the methods section to clarify this aspect of our calculations. The revised version now reads as:

   > "Instruments logged data at different intervals (1-10 seconds), and all data was averaged over 10 seconds. Freeway segments were demarcated in the 10 second data time series based on location information collected using the GPS."
   >
   > "All statistics (median or first percentile concentration) required to calculate EF using Equation 1 were determined from the time series for each freeway segment, typically tens of miles long. If multiple runs were conducted on a freeway within a day, the time series for each run was analyzed separately."

4. *The second major point which needs to be clearly brought out is that the work assumes that there are only two classes of vehicle, i.e. light-duty vehicles operating on gasoline fuel and heavy-duty vehicles operating on diesel fuel. Are things actually that clear-cut in California? Could the techniques be applied in Europe where there is a substantial light-duty fleet using diesel?*

The study does assume that light duty vehicles (LDV) are mostly gasoline driven and heavy duty vehicles (HDV) are diesel driven, but not without corrections. Good statistics (both statewide and at LA county level) are available for vehicle categories that violate our assumption. Only 0.27% of LDV vehicles are diesel powered. Of the total vehicle miles travelled (VMT) in LA county, only 10% are by HDV, which ware assumed to be diesel powered and only 0.06% of heavy HDV and 4% of all HDV (light, medium, and heavy duty) use gasoline fuel. Weighted fuel efficiency (miles per gallon) was used to accommodate these small violations of our assumptions. Complete detail on these corrections was provided in the Supporting Information as Table S.3, but the following sentence has been added to the manuscript to draw attention to the details in the Supporting Information.

> "More details are given in Section S.2 in SI, where corrections applied to fuel efficiency to meet the studies' assumptions have also been detailed (Table S3)."

The term "LDV" has been replaced by "LDG," i.e., light duty gasoline and the term "HDV" has been replaced by "HDD," i.e., heavy duty diesel, which is more accurate representation of the study's assumptions.

On the issue of applicability in Europe**,** as long as reliable statistics on the fuel type division are available the same methodology of partitioning emissions could be extended to three groups (HDV – diesel, LDV – diesel and LDV – gasoline). Though mathematically feasible, we think it will be challenging to obtain the fuel type statistics required. The technique requires at least two of the three $\Delta P/\Delta CO_2$ ratios to be known for direct application. These could be obtained from roadway measurements in areas where only vehicles with certain fuel type are allowed (if they exist in Europe, though authors are not as familiar with European roadways). It is unlikely that there are places where only gasoline or diesel LDV are allowed, ruling out the possibility for two known $\Delta P/\Delta CO_2$ ratios. But it is probably possible to find locations near ports/railway yards that are heavily dominated by heavy duty diesel truck traffic. Once at least one, i.e., $\Delta P/\Delta CO_2$ ratios for diesel HDV are known, our suggestion is to first apportion diesel HDV emissions out from mixed fuel traffic measurements as opposed to our approach where we had access to only gasoline LDV traffic emission and we apportioned those out using the measured ratios. The remaining emission will then have to be apportioned out in diesel LDV and gasoline LDV fractions, but since both $\Delta P/\Delta CO_2$ ratios are unknowns, they will have to be estimated. Each observation post HDD apportionment would give a set of linear equations (of the type below for each observation) which could be solved to get an approximate solution through linear least squares for an over-determined linear system.

Equation for each observation, where 'o' is number of observations:

$(\Delta P/\Delta CO2)\_LDG * \text{fraction of } LDG_o + (\Delta P/\Delta CO2)\_LDD * \text{fraction of } LDD_o = (\Delta P/\Delta CO2)\_apportioned_{LD}$

However, the accuracy of regressed LDG and LDD ΔP/ΔCO2 ratios would depend on the accuracy of fraction of LDG and LDD input, which is likely to be less accurately known than the division between LD and HD vehicles. If fraction of LDG and LDD estimates is very uncertain, it might be a better approach to chase individual vehicles and calculate individual EFs.

Other points which should be addressed are as follows:

5. *Page 18719, sentence starting on last line – this states that "our study used a hybrid approach, combining individual plume impacts into longer averages that still manage to capture the spread and skew of individual EFs". This is an unsupported statement that needs to be backed up by quantitative information as suggested above.*

In our hybrid approach, where short time intervals rather than individual plumes are analyzed, we were able to capture the variability and distribution of emission factors (EFs), while vastly simplifying the analysis requited to make individual plume EF determinations. The spread of EFs in this study matches the spread reported by the latest study in California (now added to the manuscript) that measured individual heavy duty vehicles (HDV) (Dallmann et al., 2012; ES&T). This was one of the few studies of individual plume EFs that reflected normal power operation, in contrast to the majority of other studies that were conducted in locations of high acceleration or roadway grade. This adds additional confirmation that we were able to capture the spread of EFs on the basis of comparison between our results and those from other studies that target individual vehicles.

We have also edited the title of the paper. It is now more appropriately titled as "Efficient determination of vehicle emission factors…" instead of "Cost effective determination of vehicle ……"

Also, regarding "individual vehicle EFs," only dynamometer-based measurements are capable of fully capturing inter-vehicle variability, since all driving conditions are controlled. All other studies (remote sensing, individual vehicles with stationary or mobile set-up) provide a 'snapshot' of instantaneous EFs for a vehicle. Often, these instantaneous EFs across vehicles are compared and conclusions on inter-vehicle variability are drawn where in fact EFs for different vehicles under different driving mode are being compared. It is difficult to draw such conclusions without knowing the full extent of variation within an individual vehicle for different driving conditions.

6. *Page 18725, line 7 – what is NOx/NO? Does this mean a ratio, or both NOx and NO, or what? Page 18726 – line 15 – this refers to a NO2/NOx fraction but seems to have a different meaning from NOx/NO referred to above.*

In the revised version, NOx/NO has been substituted with NOx and NO wherever it refers to something that applies to both species.

7. *Page 18727 – the work on diurnal variations in emission rates is interesting but inadequately explained. Are the diurnal variations based wholly on the measured vehicle counts or do they take account of the speed and driving mode dependence of the emission factors?*

We did not resolve our EFs by speed; we used study average EF to compute the total freeway emission rates (ER) but did take into account diurnal variations in vehicle miles traveled (VMT). It is possible to measure the EF variation with speed using our mobile sampling technique, but this was beyond the scope of this study. The point we were trying to illustrate by the ER calculation was that the conventionally-held belief that LA freeways with higher HDD traffic are significantly worse sources of pollution needs reconsideration. By ignoring effects of speed on EFs, we likely understated emissions during rush hours where speed is most reduced and stop-and-go conditions occur, but these diurnal differences in speed in Los Angeles are very similar across freeways, so our comparisons between freeways are still valid.

Please also see our response to Comment 3 by Referee #1.

8. *Table 1 – the labeling of footnotes appears to be incorrect especially with regard to numbers 2 and 3.*

This has been corrected in the revised version.